

Word Structure in Malayic Languages of Sumatra: Evidence from Special Registers

David Gil

Max Planck Institute for Evolutionary Anthropology, Leipzig

Abstract

The Malayic languages of Central Sumatra are amongst the world's most highly isolating languages, with very little inflectional or derivational morphology. Accordingly, there is also very little direct grammatical evidence for word boundaries and word-internal structure. What little evidence there is derives largely from a variety of "para-grammatical" phenomena involving special registers or genres, such as SMSs (text messages), pantuns (a traditional poetic form), and ludlings (secret languages).

Languages such as Minangkabau, Siak Malay and Riau Indonesian are argued to have word-internal structure along the lines presented in Figure 1 below. In accordance with this structure, an outer word contains an inner word plus optional additional material, while an inner word consists of a core foot plus optional pre-core. Evidence for the above word structure is obtained from four different sources involving special registers, as summarized in Table 1 below. Data and analyses supporting the proposed word structure are presented in the following pages.

Data

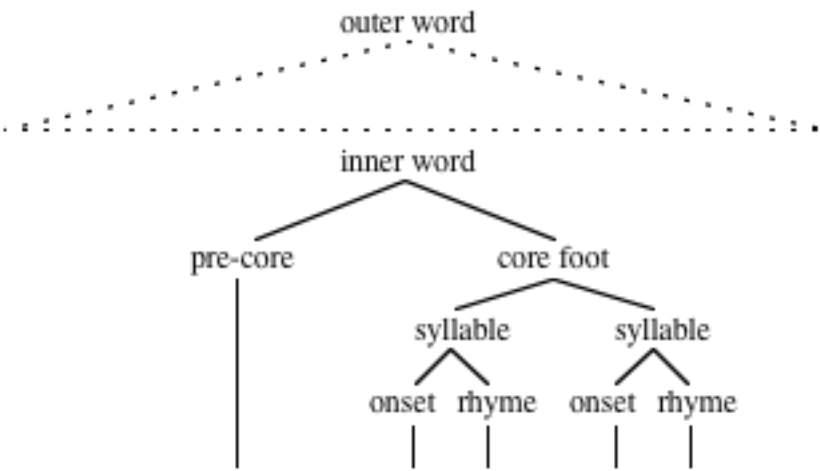


Figure 1: Word Structure

	CORE FOOT	INNER WORD	OUTER WORD
spelling		X	X
pantun rhythm		X	
Warasa ludling	X		X
Sabaha ludling		X	

Table 1: Sources of Evidence for Word Structure Categories

	<i>has own core foot</i>	<i>occurs within inner word</i>
<i>stem</i>	√	√
<i>clitic</i>	-	-
<i>affix</i>	-	√

Table 2: Three Basic Kinds of Morphemes

<i>Morphological Category</i>	<i>Syntactic Category</i>	<i>Phonological Form</i>	<i>Examples</i>
-	particle	suprasegmental	<i>lah</i> 'CONTR', <i>do</i> 'NEG.POL', <i>'kan</i> 'Q', <i>ek</i> 'Q'
stem	S	polysyllabic	<i>makan</i> 'eat', <i>ayam</i> 'chicken', <i>merah</i> 'red', <i>aku</i> '1SG', <i>tiga</i> 'three', <i>udah</i> 'PFCT', <i>tadi</i> 'PST:PROX', <i>boleh</i> 'can', <i>sama</i> 'together', <i>tidak</i> 'NEG' [and most other morphemes]
stem	S	monosyllabic	<i>mi</i> 'noodles', <i>cat</i> 'paint', <i>cas</i> 'charge', <i>lap</i> 'wipe', <i>tes</i> 'test', <i>om</i> 'uncle', <i>rem</i> 'brake', <i>tas</i> 'bag' [and a handful of other morphemes]
stem	S/S	polysyllabic	<i>dengan</i> 'with', <i>dari</i> 'from', <i>aja</i> 'just', <i>untuk</i> 'for', <i>supaya</i> 'in order that', <i>tentang</i> 'about', <i>pada</i> 'COMP', <i>kalau</i> 'TOP', <i>tiap</i> 'every', <i>atau</i> 'or', <i>pula</i> 'CONJ.OP', <i>juga</i> 'CONJ.OP'
stem	S/S	monosyllabic	<i>yang</i> 'REL', <i>pun</i> 'CONJ.OP'
<i>Table 3: Classification of Morphemes in Riau Indonesian</i>			

<i>Morphological Category</i>	<i>Syntactic Category</i>	<i>Phonological Form</i>	<i>Examples</i>
stem/clitic	S	polysyllabic	*
stem/clitic	S	monosyllabic	<i>tak-</i> 'NEG', <i>-ni</i> 'DEM:PROX', <i>-tu</i> 'DEM:DIST'
stem/clitic	S/S	polysyllabic	*
stem/clitic	S/S	monosyllabic	<i>-nya</i> 'ASSOC'
clitic	S	polysyllabic	*
clitic	S	monosyllabic	*
clitic	S/S	polysyllabic	*
clitic	S/S	monosyllabic	<i>N-</i> [<i>me-</i>] 'AG', <i>di-</i> 'PAT', <i>ter-</i> 'NON.AG', <i>-kan</i> 'EP', <i>-in</i> 'EP'
clitic/affix	S	polysyllabic	*
clitic/affix	S	monosyllabic	*
clitic/affix	S/S	polysyllabic	*
clitic/affix	S/S	monosyllabic	<i>ber-</i> 'NON.PAT', <i>ke-</i> 'DIR', <i>ke-</i> 'ABSTR', <i>i-</i> 'DEM', <i>s-</i> 'LOC', <i>g-</i> 'like', <i>-an</i> 'AUG'
affix	S	polysyllabic	*
affix	S	monosyllabic	*
affix	S/S	polysyllabic	*
affix	S/S	monosyllabic	<i>N-</i> [most allomorphs] 'AG', <i>(s)i-</i> 'PERS', <i>se-</i> 'one'
<i>Table 3 [cont]: Classification of Morphemes</i>			

SMS Messages: Spelling

- (1) Kenapa tak ajak teman dapit, aku mau kenalan ama teman2 mu, eh dapit disana perlu ditemani nggak

Kenapa tak ajak teman David? Aku mau kenalan ama
 why NEG invite companion David 1:SG want recognize-AUG together
 teman-teman mu Eh David di sana perlu ditemani nggak?
 DISTR~companion POSS:2 EXCL David LOC there need PAT-companion-EP NEG
 'Why don't you invite your friends? I want to get to know your friends. Do you need a companion there?'
- (2) Sayang maaf ya,sur baru sekarang membalas sms nya habisnya hp sur di tarok di lemari jadi ngak,i miss you.

Sayang maaf ya, Sur baru sekarang membalas smsnya habisnya
 compassion forgive yes Sur new now AG-reply sms-ASSOC finish-ASSOC
 hp sur ditarok di lemari, jadi nggak? I miss you
 handphone Sur PAT-put LOC cupboard become NEG
 'Sorry, love, I'm only just responding to your message, because I put the handphone in the cupboard, so will it happen? I miss you'
- (3) buah salak di tepi perigi jangan sampai jatuh sebiji kenapa david rindu kan pii spai2 lupakan adi

Buah salak di tepi perigi, Jangan sampai jatuh sebiji,
 fruit salak LOC edge well NEG.IMP reach fall one-CLF
 Kenapa David rindukan Pii, Sampai-sampai lupakan Adi
 why David miss-EP Pii DISTR~reach forget-EP Adi
 'Salak fruit on the edge of the well, Do not let one happen to fall, Why is it you miss Pii, so much so that you forget Adi'
- (4) Pit anton udah datang ke pakning,dapit datang kepakning kapan pit

Vid, Anton udah datang ke Pakning, David datang ke Pakningkapan, Vid
 FAM|David Anton PFCT arrive DIR Pakning David arrive DIR Pakningwhen FAM|David
 'David, Anton's already arrived in Pakning, when are you arriving, David?'
- (5) Udah puas aku ngirim tapi takmasuk2 juga, entah apa sebab nya tak tau lah dapit

Udah puas aku ngirim tapi tak masuk-masuk juga,
 PFCT satiated 1:SG AG-send but NEG DISTR~enter also
 entah apa sebabnya, tak tau lah, David
 uncertain what reason-ASSOCNEG know CONTR David
 'I've already had enough of my messages not getting through, I don't know what the reason is, I just don't know, David'

<i>form</i>	<i>standard orthography</i>	<i>classification (as per Table 3)</i>	<i>total tokens</i>	<i>together</i>	<i>separate</i>	<i>% separate</i>
<i>tu</i>	word	stem/clitic	45	0	45	100
<i>yang</i>	word	stem	267	1	266	<100
<i>nggak</i>	word	stem	214	2	212	99
<i>dari</i>	word	stem	245	6	239	98
<i>tak</i>	word	stem/clitic	758	39	719	95
<i>ni</i>	word	stem/clitic	144	9	135	94
<i>pun</i>	word	stem	46	4	42	91
<i>di</i> PASS	-	clitic	221	51	170	77
<i>di</i> LOC	word	clitic	597	181	416	70
<i>nya</i>	-	stem/clitic	569	170	399	70
<i>ke</i> DIR	word	clitic/affix	309	133	176	57
<i>kan</i>	-	clitic	236	109	127	54
<i>si</i>	word	affix	17	8	9	53
[other] <i>in</i>	-	clitic	53	50	3	6
<i>ter</i>	-	clitic	81	79	2	2
<i>ber</i>	-	clitic/affix	209	207	2	1
(<i>ngapa</i>) <i>in</i>	-	clitic	249	248	1	>0
<i>se</i>	-	affix	217	217	0	0

*Table 4: Word Separation in the SMS Corpus (out of 4142 utterances)
Ranked by % Separate: All*

<i>form</i>	<i>standard orthography</i>	<i>classification (as per Table 3)</i>	<i>total tokens</i>	<i>together</i>	<i>separate</i>	<i>% separate</i>
<i>yang</i>	word	stem	267	1	266	<100
<i>nggak</i>	word	stem	214	2	212	99
<i>dari</i>	word	stem	245	6	239	98
<i>pun</i>	word	stem	46	4	42	91
<i>tu</i>	word	stem/clitic	45	0	45	100
<i>tak</i>	word	stem/clitic	758	39	719	95
<i>ni</i>	word	stem/clitic	144	9	135	94
<i>nya</i>	-	stem/clitic	569	170	399	70
<i>di</i> PASS	-	clitic	221	51	170	77
<i>di</i> LOC	word	clitic	597	181	416	70
<i>kan</i>	-	clitic	236	109	127	54
[other] <i>in</i>	-	clitic	53	50	3	6
<i>ter</i>	-	clitic	81	79	2	2
<i>(ngapa)</i> <i>in</i>	-	clitic	249	248	1	>0
<i>ke</i> DIR	word	clitic/affix	309	133	176	57
<i>ber</i>	-	clitic/affix	209	207	2	1
<i>si</i>	word	affix	17	8	9	53
<i>se</i>	-	affix	217	217	0	0

Table 5: Word Separation in the SMS Corpus (out of 4142 utterances)
Ranked by % Separate: Within Structural Types

<i>form</i>	<i>standard orthography</i>	<i>classification (as per Table 3)</i>	<i>total tokens</i>	<i>together</i>	<i>separate</i>	<i>% separate</i>
<i>tu</i>	word	stem/clitic	45	0	45	100
<i>yang</i>	word	stem	267	1	266	<100
<i>nggak</i>	word	stem	214	2	212	99
<i>dari</i>	word	stem	245	6	239	98
<i>tak</i>	word	stem/clitic	758	39	719	95
<i>ni</i>	word	stem/clitic	144	9	135	94
<i>pun</i>	word	stem	46	4	42	91
<i>di</i> LOC	word	clitic	597	181	416	70
<i>ke</i> DIR	word	clitic/affix	309	133	176	57
<i>si</i>	word	affix	17	8	9	53
<i>di</i> PASS	-	clitic	221	51	170	77
<i>nya</i>	-	stem/clitic	569	170	399	70
<i>kan</i>	-	clitic	236	109	127	54
[other] <i>in</i>	-	clitic	53	50	3	6
<i>ter</i>	-	clitic	81	79	2	2
<i>ber</i>	-	clitic/affix	209	207	2	1
(<i>ngapa</i>) <i>in</i>	-	clitic	249	248	1	>0
<i>se</i>	-	affix	217	217	0	0

*Table 6: Word Separation in the SMS Corpus (out of 4142 utterances)
Ranked by % Separate: Within Orthographic Types*

Pantun: Rhythm

- | | | | | |
|-----|--|--|---|--|
| (6) | Bukan
NEG

Dulang
tray

Bukan
NEG

Datang
arrive | dulang
tray

kami
1

datang
arrive

kami
1 | sembarang
one-AG-thing

dulang
tray

sembarang
one-AG-thing

nak
want | dulang
tray

kuning
yellow

datang
arrive

merunding
AG-discuss |
|-----|--|--|---|--|

'The tray is not just any tray
 Our tray is a yellow tray
 The arrival is not just any arrival
 Our arrival is to negotiate a wedding'

- | | | | | |
|-----|--|--|--|---|
| (7) | Hari
day

Hari
day

Hari
day

Hari
day | ini
DEM-DEM.PROX

besok
tomorrow

ini
DEM-DEM.PROX

besok
tomorrow | menanam
AG-plant

menanam
AG-plant

kita
1.2

kita
1.2 | tebu
corn

serai
lemongrass

bertemu
NON.PAT-meet

bercerai
NON.PAT-divorce |
|-----|--|--|--|---|

'Today planting corn
 Tomorrow planting lemongrass
 Today we meet
 Tomorrow we divorce'

- | | | | | |
|-----|---|---|--|--|
| (8) | <i>Jalan-</i>
DISTR~

Nampak
see

Biar
let

Asal
reason | <i>jalan</i>
go

rumah
house

mati
die

dapat
get | ke kota
DIR city

berbaris-
NON.PAT-DISTR~

di ujung
PAT end

cewek
chick | Bengkalis
Bengkalis

<i>baris</i>
row

keris
dagger

yang manis
REL sweet |
|-----|---|---|--|--|

'Going to Bengkalis city
 Seeing houses all in rows
 Let one die at the tip of a dagger
 As long as one gets a sweet girl'

(9)	<i>Jalan-</i> DISTR~	<i>jalan</i> go	ke Pekan DIR Pekan...	baru ...baru
	Jangan NEG.IMP	lupa forget	membeli AG-buy	rambutan rambutan
	Sungguh true	enak nice	pengantin AG-bride(groom)	baru new
	Dalam inside	kamar room	<i>cubit-</i> DISTR~	<i>cubitan</i> pinch-AUG

'Going to Pekanbaru
Don't forget to buy rambutan
How nice are the newlyweds
Pinching each other in the room'

(10)	<i>Jalan-</i> DISTR~	<i>jalan</i> go	ke rumah DIR house	Rudy Rudy
	Singgga stop.over	sebentar one-while	kat <i>Sungai</i> near Sungai	<i>Apit</i> Apit
	David David	datang arrive	melenggang- AG-DISTR~	<i>lenggok</i> sway:ASSOC
	Nampak see	Susi Susi	langsung [x3] straight	terdiam NON.AG-quiet

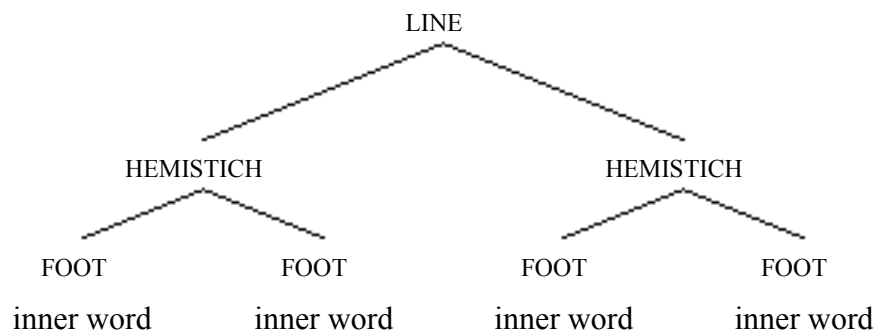
'Going to Rudy's house
Stopping over at Sungai Apit
David arrives all topsy turvy
Seeing Susi immediately falling quiet'

(11)	<i>Jalan-</i> DISTR~	<i>jalan</i> go	ke <i>Kuala</i> DIR Kuala	<i>Deli</i> Deli
	[a] Lewat pass	jalan road	berliku- NON.PAT-DISTR~	<i>liku</i> curve
	Kalau TOP	anda 2	temannya friend-ASSOC	Elly Elly
	Singgga stop.over	dulu first	di <i>Wisma</i> PAT Wisma	<i>Ratu</i> Ratu

'Going to Kuala Deli
Along a road full of curves
If you want to be friends with Elly
Stop off first at Wisma Ratu'

(12) <i>Jalan-</i>	<i>jalan</i>	ke rumah	Yendy
DISTR~	go	DIR house	Yendy
Singga	sebentar	memakan	kueh
stop.over	one-while	AG-eat	cake
Apa	buat	Yendy	ke belakang?
what	do	Yendy	DIR back
Main	cewek	dengan	Geta
play	chick	with	Geta
'Going to Yendy's house			
Stopping over to eat some cake			
What's Yendy doing out the back?			
Playing around with Geta'			

(13) Kalau	anda	ingin	melayang
TOP	2	desire	AG-kite
Jangan	lupa	membawa	banci
NEG.IMP	forget	AG-carry	axe
David	datang	ingin	belajar
David	arrive	desire	NON.PAT-learn
belajar	bahasa	<i>Sungai</i>	<i>Selari</i>
NON.PAT-learn	language	Sungai	Selari
'If you want to fly a kite			
Don't forget to bring an axe			
David has come and wants to learn			
To learn the language of Sungai Selari'			

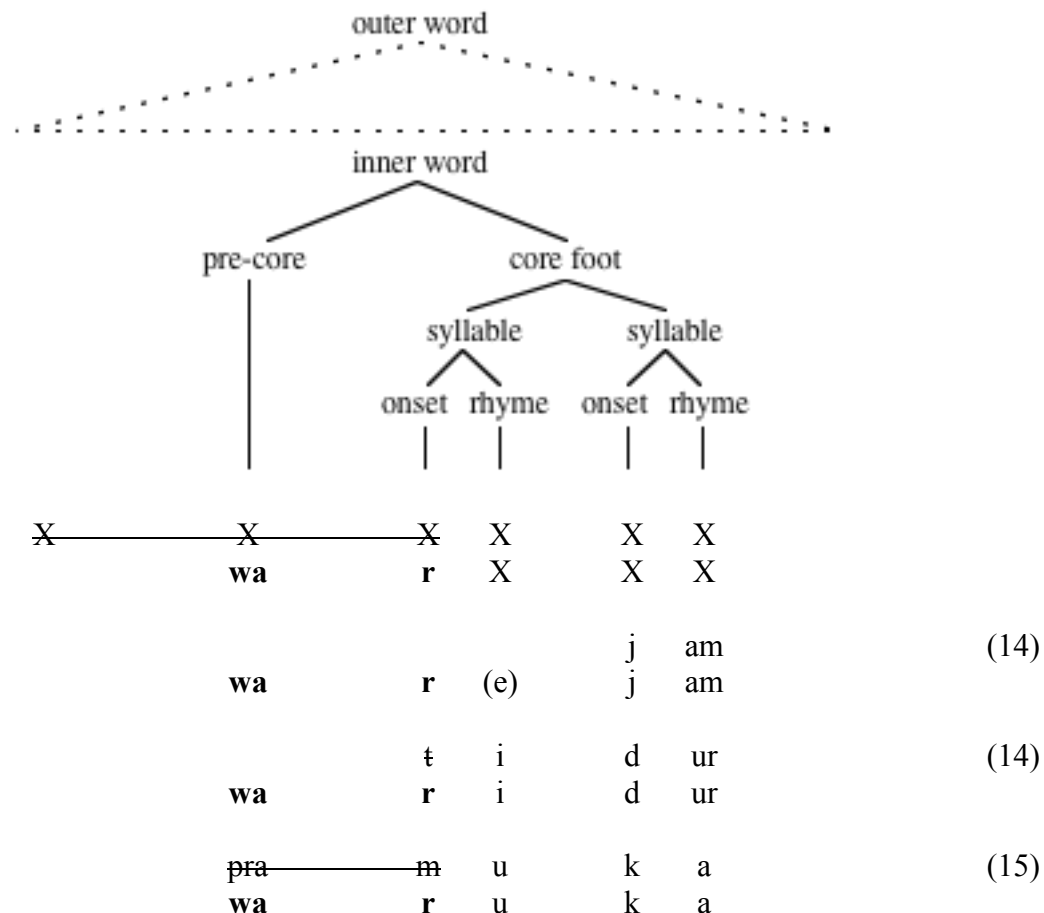


Warasa Ludling

- (14) Waraku, wara- waraku warendak *waridur* do, waraku wara- warari
 Aku, aku tendak tidur do, aku hari
 1:SG 1:SG NEG:want sleep NEG.POL 1:SG day
warejam warailan waraen warai- walai- *warelai* *wareson*
 jam sembilan maen Play Station
 hour nine play Play Station
 [Discussing plans]
 'I don't want to sleep, I play Play Station at nine o'clock'
- (15) W- warabtu warajo, waraku *warisa* waringgu warami waruo warowo *waruka*
 Sabtu ajo, aku tak bisa minggu kami duo Bowo pramuka
 Saturday just 1:SG NEG can Sunday 1 two Bowo scouts
 [Discussing planned beach trip, arguing whether to go on Saturday or Sunday]
 'Make it Saturday, I can't go on Sunday, me and Bowo have scouts'
- (16) *Waripai*, warengkau warapo waranyak warawat?
 I Pai, engkau kenapa banyak jerawat?
 PERS Paisal|FAM 2 why lots pimple
 [Small talk]
 'Paisal, why do you have so many pimples?'
- (17) Waravid *warajok*
 David merajok
 David AG-sulk
 [Discussing planned beach trip; my silence,intended to avoid having my voice be part of
 the recording, is misinterpreted]
 'David's sulking'
- (18) Waravid ek warengkau *warangonkan* waraku lah
 David ek engkau bangonkan aku lah
 David Q 2 get.up-EP 1:SG CONTR
 [Discussing planned beach trip]
 'David, you wake me up'
- (19) Dadanyak dadetul *dadulu* nyo ek
 Banyak betul bulu nyo ek
 lots really feather ASSOC Q
 [About me]
 'He's got lots of body hair, hasn't he?'

- (20) Wariam, *warowok ni* warau waridur
 Diam, Bowok ni mau tidur
 quiet Bowok DEM:PROX want sleep
 [Berating his friends for making too much noise]
 'Quiet, Bowok wants to sleep'

- (21) Warami waredap *Warelat Wararu* waruduk warano waredap warisap
 Kami sedap Selat Baru duduk sano sedap hisap
 I nice Selat Baru sit there nice suck
 a warinom-warinom *warakan-warakan* ek, waravid ek
 a minom-minom makan-makan ek David ek
 EXCL DISTR~drink DISTR~eat Q David Q
 [Discussing planned beach trip to Selat Baru]
 'I like it in Selat Baru, sitting around there, it's nice, smoking, drinking, eating, right David'



me ————— wa r i p ai (16)

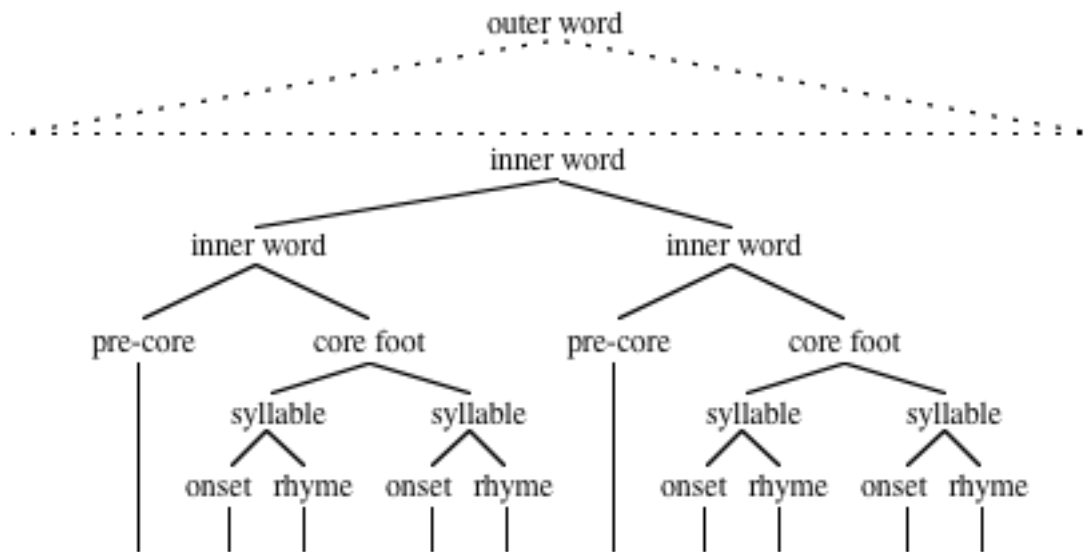
wa r a j ok (17)

tak ————— wa r i s a (15)

wa b a ng on kan (18)

da b u l u nyo (19)

wa B o w ok ni (20)



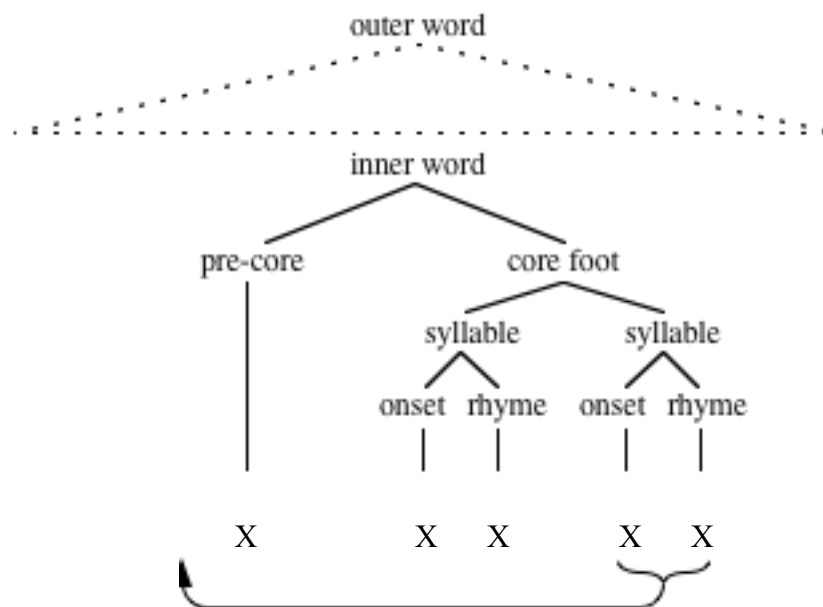
wa p e l ai (14)

wa S e l at (21)

wa m a k an (21)

Sabaha Ludling

- (22) *Diroahsunyo* likba *japke*
 Disuruahnyo balik kejak
 PAT-order-ASSOC return wink
 [Speaker worried he should be back at work]
 'He said we should return for a bit'
- (23) *Tibara* yang *dangsaga* tui yang waka hekli tui
 Barati yang sagadang itu yang awak lihek itu
 NON.PAT-meaning REL one-big DEM-DEM:DIST REL 1.2 see DEM-DEM:DIST
 [Discussing different kinds of fish]
 'So the ones that are as big as that, the ones that we saw'
- (24) Wekce angwak *Gutsi* 'kan
 Cewek wakang si Gut 'kan
 [Teasing friend]
 'Your girlfriend is Gut'
- (25) *Molanyo* *jokara* Ja- Jakarta Ndaki?
 Lamonyo karajo Indak?
 long.time-ASSOC work Jakarta NEG
 [Asking me about my travel plans]
 'Are you going to be working in Jakarta for a long time?'
- (26) *Lakga-lakga* mika
 Galak-galak kami
 DISTR~laugh 1
 [From a long narrative]
 'We all laughed'
- (27) Doa, dangga *kani* *hanlounyo*
 Ado gadang ikan louhannyo
 exist big fish lou han
 [Arguing whether a certain place has lou han fish]
 'Yes there are, there are big lou han fish'



k e j ap (22)
j ap k e

ka r a j o (25)
jo k a r a

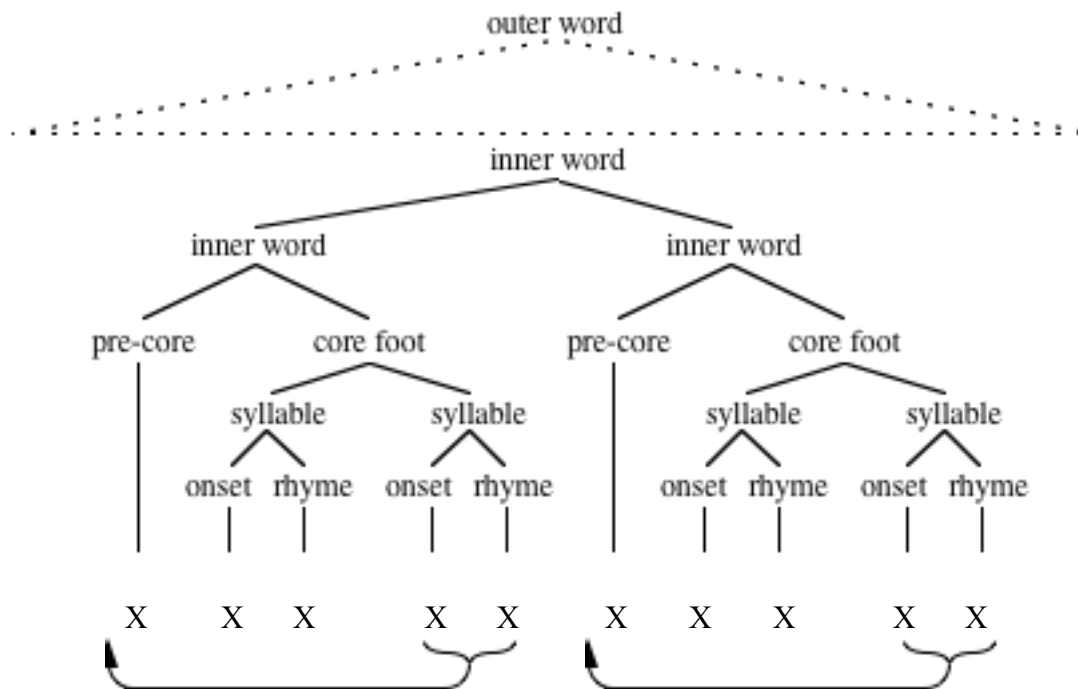
s i g ut (24)
g ut s i

ba r a t i (23)
ti b a r a

sa g a d ang (23)
dang s a g a

l a m o nyo (25)
m o l a nyo

di s u r uah nyo (22)
di r oah s u nyo



g a l ak (26)
l ak g a

i k an
k an i nyo (27)
h ou l ou